A Complete 3-D Canonical Piecewise-Linear Representation

Chengtao Wen, Shuning Wang
Institute of Systems Engineering, Department of Automation,Tsinghua University
Beijing, 100084, P. R. China
wct01@mails.tsinghua.edu.cn

***Abstract* - A complete 3-D Canonical Piecewise-Linear (CPWL) representation is developed constructively in this paper. The key to the representation is the establishment of the explicit functional formulation of basis function. It is proved that basis function is the most elementary generating function from which a fully general 3-D PWL function can be formulated. This CPWL representation laid a solid theoretical foundation for the development of a general nonlinear approximation, which can be seen as an extended version of the well-established Hinging Hyperplane Algorithm.**

***Index Terms* - Piecewise-linear representation, basis function, Piecewise-linear modeling, nonlinear approximation.**

I. INTRODUCTION

The continuous Piecewise-Linear (PWL) function is a promising tool in many fields where strong nonlinearity exits such as nonlinear system modeling, simulation and control. It prevents the PWL function's wider application that conventional representation brings in too many parameters. Then the compact representation is of especially importance because it can compress local linear functions and region boundaries into one global analytic form and squeezes out most of the redundant parameters.

The first step in this direction is the canonical representation proposed by Chua and Kang [1] and will be referred to as Chua1. The model can express any PWL function possessing the consistent variation property. Three extensions have been presented in [2-4] respectively, both of which are capable of expressing any 2-D PWL function. In [5] it was proved that any PWL function of *n* variables could be represented by no more than *n*-level nested absolute-value functions. By detailed analysis of the structure of the domain space, a canonical PWL representation is put up in [6]. In this model, the minimal degenerate intersection is defined that constitutes the "building block", from which a fully general PWL function can be formulated. Using the theoretical framework in [6], a concrete functional form is proposed for PWL functions defined over a minimal degenerate intersection [7].

Originated from the lattice PWL representation [8], a new multi-level nested absolute-value representation is built up constructively by Wang [9] and will be denoted as Wang's model. Since the most elementary "building block" is not given in this model, it is not the simplest form of canonical representation.

Although much effort has been done, it is still an open problem to find a practical canonical representation for PWL functions even in 3-D space. Very often the greatest difficulty in extending a result to *n*-dimensions is encountered in going from two to three dimensions since our geometrical intuition is more reliable with the ability to draw pictures in 2-D space.

The main purpose of the paper is to establish a novel canonical representation for all PWL functions with three variables. The representation capability of the new model is proved constructively in the form of an Algebraic Cutting Algorithm. Finally, since this CPWL presentation can be seen as a natural continuation to the Hinging Hyperplane Algorithm [10], it can also approximate any continuous nonlinear function to an arbitrary precision. This lays a solid theoretical foundation for the development of a general nonlinear approximation algorithm.

II. SIMPLIFICATION OF WANG'S MODEL

The main result of [9] can be summarized in the following lemma.

Lemma 1: For any PWL function $p: R^n \to R, n \geq 2$, there always exist a positive $m \in Z^+$, a group of real numbers $C_i \in R, i \in Z(m)$ and a set of continuous PWL functions $d_i: R^{n-1} \to R, 0 \leq i \leq m-1$ such that:

$$p(x) = d_0(\hat{x}) + c_1 x_n + \sum_{i=2}^{m} c_i \rho(x_n - d_{i-1}(\hat{x})) \tag{1}$$

where $x = [\hat{x}^T \quad x_n]^T \in R^n, \hat{x} \in R^{n-1}$ and $\rho(x) = \max(0, x)$

The significance of Wang's model is that it discovers the latent relationships between the PWL functions defined in neighboring dimensions.

Definition 1: A continuous 2-D PWL function *B* is defined as 2-D basis function if it takes one of the following four forms

$$\begin{aligned} &\max(a^T x, \max(a_1^T x, a_2^T x)) \\ &\max(a^T x, \min(a_1^T x, a_2^T x)) \\ &\min(a^T x, \min(a_1^T x, a_2^T x)) \\ &\min(a^T x, \max(a_1^T x, a_2^T x)) \end{aligned} \tag{2}$$

where $a, a_1, a_2 \in R^3$.

Here Li's Geometrical Cutting Algorithm [4] is formulated algebraically in Lemma 2.

Lemma 2: Assume that P_m is a 2-D PWL function with m regions[*]. Assume further that R_1, R_2 are two adjacent regions and the local functions are $l_i(x) = a_i^T x, a_i \in R^2, (i = 1,2)$. Then there must be a linear function $L(x) = a^T x, a_i \in R^2$ and two complementary sections $D_1 = R_1 \cup R_2$ and $D_2 = R^2 - D_1$ such that:

$$P_1 = \begin{cases} P_m & (x_1, x_2) \in D_1 \\ a^T x & (x_1, x_2) \in D_2 \end{cases} \quad (3)$$

$$P_2 = \begin{cases} a^T x & (x_1, x_2) \in D_1 \\ P_m & (x_1, x_2) \in D_2 \end{cases} \quad (4)$$

Therefore, for any $x \in R^2$, we have

$$P_m = P_1 + P_2 - a^T x \quad (5)$$

where P_1 is the 2-D basis function defined in (2), P_2 consists of m-1 regions, which can be described as $\max(a^T x, P_{m-2})$ or $\min(a^T x, P_{m-2})$ and P_{m-2} is a PWL function with m-2 regions, each local linear function of which is identical with that of P_m.

It is demonstrated in Lemma 2 that any 2-D PWL function can be equivalently transformed into an addition of simpler functions with fewer regions until it has degenerated into a 2-D basis function.

Definition 2: A PWL function $\max(B, x_3)$ is called a 3-D hinging hyperplane if B is a 2-D basis function.

Lemma 3: Any 3-D PWL function with one intersection at the origin can be represented by a superposition of 3-D hinging hyperplanes.

Proof: Let $p(x)$ be a 3-D PWL function with one intersection at the origin. Due to (1), we have

$$p(x) = \sum_{i=1}^{m} c_i d_{i-1}(x_1, x_2) + c_1 x_3 + \sum_{i=2}^{m} c_i \max(x_3, d_{i-1}(x_1, x_2)) \quad (6)$$

Take the term of $\max(x_3, d_{i-1}(x_1, x_2))$ as an example. Suppose that $d_{i-1}(x_1, x_2)$ is composed of m regions, denoted also as P_m. Inspired by (5), we can define the following 3-D PWL function:

$$P = \max(P_1, x_3) + \max(P_2, x_3) - \max(a^T x, x_3) \quad (7)$$

By denoting

$$\begin{aligned} D_1' &= \{x \mid (x_1, x_2) \in D_1, x_3 \in R\} \\ D_2' &= \{x \mid (x_1, x_2) \in D_2, x_3 \in R\} \end{aligned} \quad (8)$$

we can obtain $R^3 = D_1' + D_2'$.

First consider the case that $x \in D_1'$, we can get

$$\begin{aligned} P &= \max(P_1, x_3) + \max(a^T x, x_3) - \max(a^T x, x_3) \\ &= \max(P_1 + P_2 - a^T x, x_3) = \max(P_m, x_3) \end{aligned} \quad (9)$$

When the other case occurs that $x \in D_2'$, we can obtain

$$\begin{aligned} P &= \max(a^T x, x_3) + \max(P_2, x_3) - \max(a^T x, x_3) \\ &= \max(P_1 + P_2 - a^T x, x_3) = \max(P_m, x_3) \end{aligned} \quad (10)$$

Accordingly, for any $(x_1, x_2, x_3) \in R^3$, we have

$$P = \max(P_m, x_3) = \max(d_2(x_1, x_2), x_3)) \quad (11)$$

By virtue of (11), we can get that $\max(P_m, x_3)$ can be decomposed into a linear combination of three simpler PWL functions, the most complicated one of which consists of m-1 regions. If $m - 1 \geq 2$, the algorithm above can be run recursively. Finally we can obtain

$$\max(P_m, x_3) = \sum_{i=1}^{k} \max(P_{1i}, x_3) - \sum_{i=1}^{k-1} a_i^T x, k \in Z^+ \quad (12)$$

Since the other terms in (6) are at most a 2-D PWL function, they all can be expressed by special kinds of 3-D Hinging Hyperplanes. Then the proof of Lemma 3 is completed here.

III. Canonical Representation of 3-D PWL Function

Definition 3: A PWL function $\max(0, l_1, l_2, l_3)$ is defined as basis function if $l_i(x) = a_i^T x, a_i \in R^3$ where the parameter vectors $a_i (i = 1,2,3)$ are linearly independent.

Lemma 4: Any 3-D hinging hyperplane can be represented by a linear combination of basis functions.

Proof: For any 3-D hinging hyperplane, one of the four cases may occur

Case 1:

$$\max(x_3, \max(L, \max(l_1, l_2))) = x_3 + \max(0, L', l_1', l_2') \quad (13)$$

Case 2: Due to the following equivalent transformation

$$\min(A, B) = A + B - \max(A, B) \quad (14)$$

we have

$$\begin{aligned} &\max(x_3, \min(L, \min(l_1, l_2))) \\ &= -\min(-x_3, \max(-L, -l_1, -l_2)) \\ &= \max(0, -L + x_3, -l_1 + x_3, -l_2 + x_3) \\ &- \max(-L, -l_1, -l_2) \end{aligned} \quad (15)$$

Case 3: Define a PWL function

$$P' = \max(A, \min(B, C)) + \max(A, B, C) \quad (16)$$

we can get

$$P' = \begin{cases} \max(A, B) + \max(A, C) & B \leq C \\ \max(A, C) + \max(A, B) & B > C \end{cases} \quad (17)$$

which leads to

$$\begin{aligned} &\max(A, \min(B, C)) \\ &= \max(A, B) + \max(A, C) - \max(A, B, C) \end{aligned} \quad (18)$$

Then for the third case, we have

$$\begin{aligned} &\max(x_3, \min(L, \max(l_1, l_2))) \\ &= \max(x_3, L) + \max(x_3, l_1, l_2) - \max(x_3, L, l_1, l_2) \end{aligned} \quad (19)$$

Case 4:

$$\begin{aligned} &\max(x_3, \max(L, \min(l_1, l_2))) \\ &= \max(\max(x_3, L), \min(l_1, l_2)) \\ &= \max(x_3, L, l_1) + \max(x_3, L, l_2) - \max(x_3, L, l_1, l_2) \end{aligned} \quad (20)$$

Finally, we can represent any 3-D hinging hyperplane H as follows:

[*] A region is a partition of the domain where the PWL function degenerates into one of its local linear functions.

$$H = \sum_{i=1}^{k1} \sigma_{1i} \max(0, l_{1i}, l_{2i}, l_{3i}) + \sum_{j=1}^{k2} \sigma_{2j} \max(0, l'_{1j}, l'_{2j}) + \sum_{k=1}^{k3} \sigma_{3k} \max(0, l'_{1k}) \quad (21)$$

where $\sigma_{ij} = \pm 1$.Here the result of Lemma 4 is proved.

Theorem 1: Any 3-D PWL function can be represented by a linear combination of basis functions.

Proof: Any PWL function with an intersection located arbitrarily in the domain space can be translated to the origin through a linear transformation by adding a constant contribution to every local linear function.

According to [6], any 2-D PWL function with multi second-order degenerate intersections can be rewritten into a superposition of PWL functions, each of which is defined over a degenerate intersection of same order. This completes the proof of Theorem 1.

IV. Basis Function

In three dimensions, any continuous PWL function defined on a third-order minimal degenerate intersection can also be further decomposed into a linear combination of basis functions, so the basis function is a much more elementary "generating function" than the third-order minimal degenerate intersection. From the geometrical viewpoint, the domain of basis function consists of four linear regions intersecting at one point. Any couple of these regions is adjacent with a plane to separate them and every group of 3 regions intersects together to form a radial as the common boundary. It follows that the boundary configuration of basis function is the simplest degenerate boundary intersection that can produce a third-order Jacobian difference function [6]. Therefore, the basis function is the simplest type of third-order minimal degenerate intersection.

Example 1:

$$P(x) = \max(\min(x_1, x_2 + x_3, x_1 + x_2 - x_3), \max(2x_1 + x_2 - x_3, x_1 + 2x_2, x_1 + x_2 + x_3)) \quad (22)$$

is a PWL function, the domain structure of which is visualized in Fig.1. The regions and corresponding local linear functions are illustrated in Table 1.

TABLE I

Regions and Local Linear Functions

Region	Local Function	Region	Local Function
$OABD$	$2x_1 + x_2 - x_3$	$OABD'$	$x_2 + x_3$
$OBCD$	$x_1 + 2x_2$	$OBCD'$	x_1
$OACD$	$x_1 + x_2 + x_3$	$OACD'$	$x_1 + x_2 - x_3$

Take the 1-D linear manifold ε_1 for an example, we can get

$$\varepsilon_1 = OAB \cap OAC \cap OAD \quad (23)$$

By definition, ε_1 is a second-order minimal degenerate intersection. Since there are four such manifolds $\{\varepsilon_i, i = 0, \cdots, 3\}$, $P(x)$ is defined over a third-order minimal degenerate intersection. Choosing $L(x) = x_1 + x_2$ as the Cutting Hyperplane, we can obtain two complementary regions in the domain space

$$D_1 = \{x \mid x_1 \geq 0, x_2 \geq 0, x_3 \geq 0; x \in R^3\}, \quad D_2 = R^3 - D_1 \quad (24)$$

If $x \in D_1$, we can get

$$\min(2x_1 + x_2 - x_3, x_1 + 2x_2, x_1 + x_2 + x_3)) \geq L(x), \quad \max(x_1, x_2 + x_3, x_1 + x_2 - x_3) \leq L(x) \quad (25)$$

If $x \in D_2$, we can obtain

$$\max(2x_1 + x_2 - x_3, x_1 + 2x_2, x_1 + x_2 + x_3)) \leq L(x), \quad \min(x_1, x_2 + x_3, x_1 + x_2 - x_3) \geq L(x) \quad (26)$$

Therefore, For any $x \in R^3$

$$\begin{aligned} P(x) &= P_1(x) + P_2(x) - L(x) \\ &= \max(L(x), 2x_1 + x_2 - x_3, x_1 + 2x_2, x_1 + x_2 + x_3)) \\ &\quad + \max(L(x), \min(x_1, x_2 + x_3, x_1 + x_2 - x_3)) - L(x) \end{aligned} \quad (27)$$

where $P_1(x)$ and $P_2(x)$ are two basis functions defined over D_1, D_2, respectively (shown in Fig. 2).

After further simplification, we have

$$P(x) = x_1 + x_2 + 2\max(0, x_2, x_3, x_1 - x_3) - \max(x_2, x_3, x_1 - x_3) \quad (28)$$

It is proposed by Julian that any PWL function defined over a kth-order minimal degenerate intersection can be written in the canonical form [7]

$$F(x) = a + \beta^T x + \sum_{j=1}^{k} \sum_{l=1}^{N_k(j)} c_{j,l} \gamma^j (a_1^{(j),l,m} \beta^{(j),l;1^T} x, \cdots a_j^{(j),l,m} \beta^{(j),l;1^T} x) \quad (29)$$

where $\gamma^k(x_1, x_2, \cdots, x_k) = \max(0, \min(x_1, x_2, \cdots x_k))$

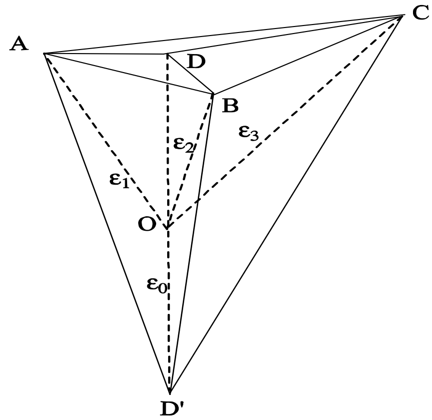

Fig.1.The domain structure of $P(x)$, which is a third-order minimal degenerate intersection.

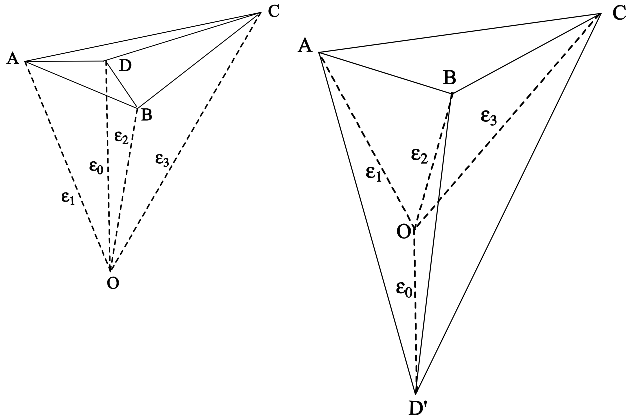

Fig.2. The domain structure of $P_1(x)$ and $P_2(x)$, both of which are 3-D basis functions.

Specifically in three-dimensional space, we have

$$\gamma^3(x_1, x_2, x_3) = \max(0, -x_1, -x_2, -x_3) - \max(-x_1, -x_2, -x_3) \quad (30)$$

Accordingly, we can rediscover the meaning of Julian's theorem as that it provides a way to represent a PWL function defined over minimal degenerate intersection with a linear combination of basis functions. Then the representation theorem in the paper can be seen as the extended version of Julian's theorem in three dimensions, where the minimal degenerate intersection is further decomposed into still simpler basis function, which is a much more elementary "building block".

In [10] Breiman developed a hinging-finding algorithm (HFA) to approximate sufficiently smooth nonlinear functions:

$$p(x) = \sum_i \delta_i \max\{a_{i1}^T x + b_{i1}, a_{i2}^T x + b_{i2}\} \quad (31)$$

where $\max\{a_{i1}^T x + b_{i1}, a_{i2}^T x + b_{i2}\}$ is called Hinging Hyperplane and $a_{ij}, b_{ij} \in R^n, \delta_i = \pm 1$.

Since the Hinging Hyperplane can be equivalently transformed into one level absolute-value function, the representation capability of (31) is essentially equal to Chua1 [11]. In two and higher dimensions, in order to achieve the required precision, HFA has to introduce much more terms to remedy the limitation of single Hinging Hyperplane. So the efficiency of the HFA is limited. Since the Hinging Hyperplane is in fact a special kind of the basis function, then this CPWL representation also has a fully general approximation capability for a continuous nonlinear function. Therefore, we are expected to propose a nonlinear approximation algorithm by substitute Hinging Hyperplanes with basis functions in the iterative scheme of HFA. The preliminary investigation shows that the new algorithm is promising.

V. CONCLUSION

For any continuous PWL function in three dimensions, a complete canonical representation model is developed, which is constructed upon basis functions instead of the third-order minimal degenerate intersections. It is proved by an Algebraic Cutting Algorithm that basis function is the most elementary "generating function". Then it can serve to be the measure of the complexity of any general 3-D PWL function.

In addition, basis function has a very simple mathematical formulation and geometrical boundary configuration. Then an efficient CPWL approximation can be developed for any continuous nonlinear function. Since it can search for the approximation function in a much bigger set of PWL function, the CPWL approximation is expected to be more efficient than the HFA.

REFERENCES

[1] Kang S. M. and Chua L. O., "A global representation of multidimensional piecewise linear functions with linear partitions," *IEEE Trans, Circuits Syst.*, vol. 25, pp. 938-940, 1978.

[2] Kahlert C. and Chua L. O., "A generalized canonical piecewise linear representation", *IEEE Trans, Circuits Syst.*, vol. 37, pp.373-382, 1990.

[3] Guzelis G. and Goknar I., "A canonical representation for piecewise affine maps and its applications to circuit analysis," *IEEE Trans, Circuits Syst.*, vol. 38, pp. 1342-1354, 1991.

[4] Li Xingye, Wang Shuning and Wang Wanbing; "A novel absolute value representation of continuous piecewise linear functions" MWSCAS 2001. Proceedings of the 44th IEEE 2001 Midwest Symposium on Circuits and Systems, Dayton, OH, 2001, Vol. 2 pp. 870-873, 2001.

[5] Lin J. -N. and Unbehauen R., 'A generalization of canonical piecewise-linear functions', *IEEE Trans, Circuits Syst.*, vol. 41, pp. 345-347, 1994.

[6] C. Kahlert and L. O. Chua, "The complete canonical piecewise-linear representation Part I: The geometry of the domain space" *IEEE Trans. Circuits Systems* , vol. 39, pp. 222-236, 1992.

[7] Julian, P.; "The complete canonical piecewise-linear representation: functional form for minimal degenerate intersections," *IEEE Trans, Circuits Syst.*, vol. 50, pp. 387-396, 2003.

[8] J. M. Tarela, E. Alonso and M. V. Martinez, "Region configurations for realizability of lattice piecewise-linear models," *Mathematical & Computer Modelling*, vol. 30 , pp 17-27, 1999.

[9] Shuning Wang, "General representation for continuous piecewise-linear functions" *IEEE Trans, Circuits Syst*, in press.

[10] Breiman,L., "Hinging hyperplanes for regression, classification and function approximation," *IEEE Trans. Inf. Theory*, vol.39, pp. 999-1013, 1993.

[11] Chengtao Wen, ShuningWang, Hao Zhang, M. J. Khan, "A novel compact piecewise-linear representation," *International Journal of Circuit Theory and Applications*, vol. 33, pp. 87-97